31 ATGGCCAACAAGCTGAGGAAATCGAACGCCATCGAATGGGCCACGGCCACCGGCACAGTACCGCTCCTGGAAAGGAGCTGCTGCCACAGC
 1 M A N K L R R S N A I E W A T A T G T V P L L E R S C C H S

91 GAGGACGCCGCACTGGAGCCCCAGCGAGCAAAACCAGCCATAGAGAACAAGCCCCCATCCTGCGCCACCTGAGCCAACTGAGCCACCTG   TM
31 E D A A L E P Q A S K T S H R E Q A P I L R H L S <u>Q L S H L</u>

181 CTCATCATCGCCGGACTGCTGATCGTCTGCTTGGCGGGCGTGACGGAGGGGCCGCCGGCATGCGCCGCTCATGTTCGAGGAGTCCGACACG
61 <u>L I I A G L L I V C L A G V</u> T E G R R H A P L M F E E S D T

271 GGCAGGCGGTCCAACCGACCAGCGGTCACCGAATGCCAGTTTTGGCAAAGTTTTGCGCGAATTGGGGTCCACCTGGTATGCGGATTTGGGT
91 G R R S N R P A V T E [C Q F G K V L R E L G S T W Y A D L G

361 CCACCCCTTCGGAGTTATGTACTGCATCAAGTGTGAATGTGTTGCGCATACCCAAGAAGCGGCGCATCGTTGCACGCGTCCAGTGTCGCAAT
121 P P F G V M Y C I K C E C V A I P K K R R I V A R V Q C R N   }CR1

451 ATCAAAAACGAGTGCCCGCCGGCCAAATGCGATGATCCCATCTCGTTGCCCGGAAAATGTGCAAGACCTGTCCCGGCGATCGAAACGAT
151 I K N E C P P A K C D D P I S L P G K C C K T C P G D R] N D
                                                                                    * *

541 ACGGATGTAGCCTTGGATGTGCCCGTGCCCAATGAAGAGGAAGAGCGCAACATGAAACATTACGCTGCGTTGCTAACGGGCCGCACCTCC
181 T D V A L D V P V P N E E E E R N M K H Y A A L L T G R T S
    *

631 TATTTCCTCAAGGGTGAGGAAATGAAGTCCATGTACACCACCTACAATCCGCAGAATCTGGTGGCCACCGCCCGTTTCCTGTTCCACAAG
211 Y F L K G E E M K S M Y T T Y N P Q N V V A T A R F L F H K

721 AAGAATCTATACTACTCCTTCTACACCTCATCGCGCAATCGGTCGTCCGCGTGCCATTCAATTCGTTGATGATGCGGGTGTAATCCTGGAG
241 K N L Y Y S F Y T S S R I G R P R A I Q F V D D A G V I L E

811 GAGCATCAACTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCACGGGCAAGATCTGA
271 E H Q L E T T L A G T L S V Y A N A T G K I STOP
                                                            * * *

FIGURE 1

```
  1  ATGGCCAACAAGCTGAGGAAATCGAACGCCATCGAATGGGCCACGGCCACCGGCACAGTACCGCTCCTGGAAAGGAGCTGCTGCCACAGC
  1  M  A  N  K  L  R  K  S  N  A  I  E  W  A  T  A  T  G  T  V  P  L  L  E  R  S  C  C  H  S

                                                                                            TM
 91  GAGGACGCCGCACTGGAGCCCCAAGCGAGCAAAACCAGCCATAGAGAACAAGCCCCCATCCTGCGCCACCTGAGCCAACTGAGCCACCTG
 31  E  D  A  A  L  E  P  Q  A  S  K  T  S  H  R  E  Q  A  P  I  L  R  H  L  S  Q  L  S  H  L

181  CTCATCATCGCCGGACTGCTGATCGTCTGCTTGGCGGGCGTGACGGAGGGGCCGCCGGCATGCGCCGCTCATGTTCGAGGAGTCCGACACG
 61  L  I  I  A  G  L  L  I  V  C  L  A  G  V  T  E  G  R  R  H  A  P  L  M  F  E  E  S  D  T

271  GGCAGGCGGTCCAACCGACCAGCGGTCACCGAATGCCAGTTTGGCAAAGTTTTGCGCGAATTGGGGTCCACCTGGTATGCGGATTTGGGT
 91  G  R  R  S  N  R  P  A  V  T  E  C  Q  F  G  K  V  L  R  E  L  G  S  T  W  Y  A  D  L  G

361  CCACCCTTCGGAGTTATGTACTGCATCAAGTGTGAATGTGTGCGCATACCCAAGAAGCGGCGCATCGTTGCACGCGTCCAGTGTCGCAAT   CR1
121  P  P  F  G  V  M  Y  C  I  K  C  E  C  V  A  I  P  K  K  R  R  I  V  A  R  V  Q  C  R  N

451  ATCAAAAACGAGTGCCCGCCGGCCAAATGCGATGATCCGATCTCGTTGCCCGGAAAATGCTGCAAGACCTGTCCCGGCGATCGAAACGAT
151  I  K  N  E  C  P  P  A  K  C  D  D  P  I  S  L  P  G  K  C  C  K  T  C  P  G  D  R  N  D

541  ACGGATGTAGCCTTGGATGTGCCCGTGCCCAATGAAGAGGAAGAGCGCAACATGAAACATTACGCTGCGTTGCTAACGGGCCGCACCTCC
181  T  D  V  A  L  D  V  P  V  P  N  E  E  E  E  R  N  M  K  H  Y  A  A  L  L  T  G  R  T  S

631  TATTTCCTCAAGGGTGAGGAAATGAAGTCCATGTACACCACCTACAATCCGCAGAATCTGGTGGCCACCGCCCGTTTCCTGTTCCACAAG
211  Y  F  L  K  G  E  E  M  K  S  M  Y  T  T  Y  N  P  Q  N  V  V  A  T  A  R  F  L  F  H  K

721  AAGAATCTATACTACTCCTTCTACACCTCATCGCGAATCGGTCGTCCGCGTGCCATTCAATTCGTTGATGATGCGGGTGTAATCCTGGAG
241  K  N  L  Y  Y  S  F  Y  T  S  S  R  I  G  R  P  R  A  I  Q  F  V  D  D  A  G  V  I  L  E

811  GAGCATCAACTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCACGGGCAAGATCGGCCGCGGCTCGAGGGTACCTCTA   pUAS
271  E  H  Q  L  E  T  T  L  A  G  T  L  S  V  Y  A  N  A  T  G  K  I  G  R  G  S  R  V  P  L    Sequences

901  GAGGATCTTTCTGAAGGAACCTTACTTCTGTCGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTCTAA
301  E  D  L  C  E  G  T  L  L  L  W  C  D  I  I  G  N  T  T  Y  R  D  L  K  L  STOP
```

FIGURE 2

```
  1  ATGGCCAACAAGCTGAGGAAATCGAACGCCATCGAATGGGCCACGGCCACCGGCACAGTACCGCTCCTGGAAAGGAGCTGCTGCCACAGC
  1   M  A  N  K  L  R  K  S  N  A  I  E  W  A  T  A  T  G  T  V  P  L  L  E  R  S  C  C  H  S

 91  GAGGACGCCGCACTGGAGCCCCAAGCGAGCAAAACCAGCCATAGAGAACAAGCCCCCATCCTGCGCCACCTGAGCCAACTGAGCCACCTG    TM
 31   E  D  A  A  L  E  P  Q  A  S  K  T  S  H  R  E  Q  A  P  I  L  R  H  L  S  Q  L  S  H  L

181  CTCATCATCGCCGGACTGCTGATCGTCTGCTTGGCGGGCGTGACGGAGGGCCGCCGGCATGCGCCGCTCATGTTCGAGGAGTCCGACACG
 61   L  I  I  A  G  L  L  I  V  C  L  A  G  V  T  E  G  R  R  H  A  P  L  M  F  E  E  S  D  T

271  GGCAGGCGGTCCAACCGACCAGCGGTCACCGAATGCCAGTTTGGCAAAGTTTTGCGCGAATTGGGGTCCACCTGGTATGCGGATTTGGGT
 91   G  R  R  S  N  R  P  A  V  T  E  | C  Q  F  G  K  V  L  R  E  L  G  S  T  W  Y  A  D  L  G

361  CCACCCTTCGGAGTTATGTACTGCATCAAGTGTGAATGTGTGCGCATACCCAAGAAGCGGCGCATCGTTGCACGCGTCCAGTGTCGCAAT   CR1:W->A
121   P  P  F  G  V  M  Y  C  I  K  C  E  C  V  A  I  P  K  K  R  R  I  V  A  R  V  Q  C  R  N

451  ATCAAAAACGAGTGCCCGCCGGCCAAATGCGATGATCCCATCTCGTTGCCCGGAAAATGCTGCAAGACCTGTCCCGGCGATCGAAACGAT
151   I  K  N  E  C  P  P  A  K  C  D  D  P  I  S  L  P  G  K  C  C  K  T  C  P  G  D  R | N  D

541  ACGGATGTAGCCTTGGATGTGCCCGTGCCCAATGAAGAGGAAGAGCGCAACATGAAACATTACGCTGCGTTGCTAACGGGCCGCACCTCC
181   T  D  V  A  L  D  V  P  V  P  N  E  E  E  E  R  N  M  K  H  Y  A  A  L  L  T  G  R  T  S

631  TATTTCCTCAAGGGTGAGGAAATGAAGTCCATGTACACCACCTACAATCCGCAGAATCTGGTGGCCACCGCCCGTTTCCTGTTCCACAAG
211   Y  F  L  K  G  E  E  M  K  S  M  Y  T  T  Y  N  P  Q  N  V  V  A  T  A  R  F  L  F  H  K

721  AAGAATCTATACTACTCCTTCTACACCTCATCGCGAATCGGTCGTCCGCGTGCCATTCAATTCGTTGATGATGCGGGTGTAATCCTGGAG
241   K  N  L  Y  Y  S  F  Y  T  S  S  R  I  G  R  P  R  A  I  Q  F  V  D  D  A  G  V  I  L  E

811  GAGCATCAACTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCAGGGGCAAGATCGGCCGCGGCTCGAGGGTACCTCTA   } pUAS
271   E  H  Q  L  E  T  T  L  A  G  T  L  S  V  Y  A  N  A  T  G  K  I | G  R  G  S  R  V  P  L |   } Sequences

901  GAGGATCTTTGTGAAGGAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTCTAA
301  | E  D  L  C  E  G  T  L  L  L  W  C  D  I  I  G  N  T  T  Y  R  D  L  K  L | STOP
```

FIGURE 3

```
  1  ATGGCCAACAAGCTGAGGAAATCGAACGCCATCGAATGGGCCACGGCCACCGGCACAGTACCGCTCCTGGAAAGGAGCTGCTGCCACAGC
  1  M  A  N  K  L  R  K  S  N  A  I  E  W  A  T  A  T  G  T  V  P  L  L  E  R  S  C  C  H  S

 91  GAGGACGCCGCACTGGAGCCCCAAGCGAGCAAAACCAGCCATAGAGAACAAGCCCCCATCCTGCGCCACCTGAGCCAACTGAGCCACCTG  TM
 31  E  D  A  A  L  E  P  Q  A  S  K  T  S  H  R  E  Q  A  P  I  L  R  H  L  S  Q  L  S  H  L

181  CTCATCATCGCCGGACTGCTGATCGTCTGCTTGGCGGGCGTGACGGAGGGCCGCCGGCATGCGCCGCTCATGTTCGAGGAGTCCGACACG
 61  L  I  I  A  G  L  L  I  V  C  L  A  G  V  T  E  G  R  R  H  A  P  L  M  F  E  E  S  D  T

271  GGCAGGCGGTCCAACCGACCAGCGGTCACCGAATGCCAGTTTGGCAAAGTTTTGCGCGAATTGGGGTCCACCTGGTATGCGGATTTGGGT
 91  G  R  R  S  N  R  P  A  V  T  E |C  Q  F  G  K  V  L  R  E  L  G  S  T  W  Y  A  D  L  G

361  CCACCCTTCGGAGTTATGTACTGCATCAAGTGTGAATGTGTGCGCATACCCAAGAAGCGGCGCATCGTTGCACGCGTCCAGTGTCGCAAT  ⎫
121  P  P  F  G  V  M  Y  C  I  K  C  E  C  V  A  I  P  K  K  R  R  I  V  A  R  V  Q  C  R  N  ⎬CR1

451  ATCAAAAACGAGTGCCCGCCGGCCAAATGCGATGATCCATCGTCGTTGCCCGGAAAATGCTGCAAGACCTGTCCCGGCGATCGAAACGAT  ⎪
151  I  K  N  E  C  P  P  A  K  C  D  D  P  S  S  L  P  G  K  C  C  K  T  C  P  G  D  R| N  D  ⎭
                                                                                      *  *

541  ACGGATGTAGCCTTGGATGTGCCCGTGCCCAATGAAGAGGAGAGCGCAACATGAAACATTACGCTGCGTTGCTAACGGGCCGCACCTCC
181  T  D  V  A  L  D  V  P  V  P  N  E  E  E  R  N  M  K  H  Y  A  A  L  L  T  G  R  T  S
     *

631  TATTTCCTCAAGGGTGAGGAAATGAAGTCCATGTACACCACCTACAATGCGCAGAATCTGGTGGCCACCGCCCGTTTCCTGTTCCACAAG
211  Y  F  L  K  G  E  E  M  K  S  M  Y  T  T  Y  N  P  Q  N  V  V  A  T  A  R  F  L  F  H  K

721  AAGAATCTATACTACTCCTTCTACACCTCATCGCGAATCGGTCGTCCGCGTGCCATTCAATTCGTTGATGATGCGGGTGTAATCCTGGAG
241  K  N  L  Y  Y  S  F  Y  T  S  S  R  I  G  R  P  R  A  I  Q  F  V  D  D  A  G  V  I  L  E

811  GAGCATCAACTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCACGGGCAAGATCGGCCGCGGCTCGAGGCAGCGCGGC  pUAS/PCR
271  E  H  Q  L  E  T  T  L  A  G  T  L  S  V  Y  A  N  A  T  G  K  I  G  R  G  S  R  N  R  G
                                                 *  *  *

901  CGCATCTTTTACCCATACGATGTTCCTGACTATGCGGGCTATCCCTATCACGTCCCGGACTATGCAGGATCCTATCCATATGACCTTCCA  ⎫
301  R  I  F |Y  P  Y  D  V  F  D  Y  A| G |Y  P  Y  D  V  P  D  Y  A| G  S |Y  P  Y  D  V  P|  ⎬SX-HA
                                                                                               ⎭IX-FLAG

991  GATTACGCTGCTCAGTGCGGCCGCGATTATACGGACGACGACGACAAATCA
331  D  Y  A| A  N  C  G  R |D  Y  K  D  D  D  D  K| STOP
```

FIGURE 4

Sequence Range: 1 to 222

```
                    10        20        30        40        50        60        70
Nog protei    MDHSQCLVTIYALMVFLGLRIDQGGCQHYLHIRPAPSENLPLVDLIEHPDPIYDPKEKDLNETLLRTLMVGHFDP

SuperSog P                                           260       270       280
[  71 ]                                              iqfvddagvIle--EhqLetTLagTLsV--yqn
                                                     ^^^^^^^^^v^v^  ^^-^^^-^^vv^^v^  ^^v
Nog protei                                           VDLIEHPDPIYDPKEKDLNETLLRTLMVGHFDP


                    90       100       110       120       130       140       150
                     *
Nog protei    ILPEERLGVEDLGELDLLLKQKPSGAMPAEIKGLEFYEGLQSKKHRLSKKLRRKLQMWLWSQTFCPVLYTWNDLG

SuperSog P            300       310   pUAS Vector
[  71 ]       IgrgsRvplEDLcEgtLLlW>
              ^v---^^v^^^^v^v-^^^^
Nog protei    ILPEERLGVEDLGELDLLLR


                   170       180       190       200       210       220
                              *
Nog protei    RYVKVGSCYSKRSCSVPEGMVCKAAKSMHLTILRWRCQRRVQQKCAWITIQYPVISECKCSC
```

FIGURE 5

```
-351  ATATTAGATGGACATGCATAATAATTATTCATGTAACTATGTGATTTTCATTTTACACGAGGTGTAAGTCAGAATTTAAAATTCTTAAAA
-261  ATTGCAATCACGGCTTATTGTACATATTTATGTATGCGATCTCATTATTATTATTATTTTGATAATATATTAGCAGCTGTTCAAA
-171  TCATCAACAAGTATAAAATTGAAAATGTTAAAAAATTGAAGCAGTCTTGGGTTTGCCTGCAACATGTTGCTGCCGATCGTTAGATGTTGCT
                                                                           20 4 cDNA start
-81   GCTGCATGTTGCCGCTGCATGTTGCCGTTGCATGTTGCCGCCGTTTGGCAACTTTATAAACACGGAGCGGATTCAGTCGTTCAGGTTCAG
10    TCGCTCTTGAATTGCGTGGGATTGCACATCGGTCGTTCGGCTTTTTGGGTTCGGCATTTAGAGAGATACGATACGATCCGATACGATCCG
100   ATCCAGTACAAAAATCAAATTCAAACCGCACTCCCGATCCGGTCGCCATCATATACACGGCGGCTCGCACCGCTATCTAGTAGATAA
190   AAAGTCGCGGAGAACGCAGTACGCATATAGCGAAACGCCAAAAAAATAAAAAAAAAAAGTCGCGTTGTCGTAATCCCTCCTCATACGAGAT
280   CGACTCTATTTTCCAGAGCAAGCTAAACACACTAGTGCTAAACCATAACTATATATCTAACTAAGGAAAACAAAGTCTCGAAACTGAAAA
370   CGAAAAGCGCAAATTTATGCAGCCGCTAAATAAAAACAGAAACCAAAACATAAAACACACTATACAAATCATACAAAACAAAAACAGCGA
460   ATCAAATAGTATAAAAAAAAAATAAATAAATGAGGAATAAAAAAACAGGCAAATAGAATAAATCCAATAAATCGGCGCGCGAAACTCGCG
550   TGTGTTATCTAATCTGCAAGAGAAGTACAAGAATCGGGTATAGAATCGGCTCTATACTATATCTATACACCTGATAGTATCTATATCCATT
640   GTGTGTGCCAGTGTGTGCGTGGCGACCTTTGTTTTTATATATTTTTTGTTGTTGTTCATACTGTGAAACGTGCTTTTTACAAGCCGGTCG
730   TTCAAAATACAAAATACTACAAATCAAATCATATACACATACATATCAGTAAAAACAAAACAAAAAACACATAAACATGGCCAAC
1                                                                                    M  A  N

820   AAGCTGAGGAAATCGAACGCCATCGAATGGGCCACGGCCACCGGCACAGTACCGCTCCTGGAAAGGAGCTGCTGCCACAGCGAGGACGCC
4     K  L  R  K  S  N  A  I  E  W  A  T  A  T  G  T  V  P  L  L  E  R  S  C  C  H  S  E  D  A

910   GCACTGGAGCCCCAAGCGAGCAAAACCAGCCATAGAGAACAAGCCCCCATCCTGCGCCACCTGAGCCAACTGAGCCACCTGCTCATCATC   TM
34    A  L  E  P  Q  A  S  K  T  S  H  R  E  Q  A  P  I  L  R  H  L  S  Q  L  S  H  L  L  I  I

1000  GCCGGACTGCTGATCGTCTGCTTGGCGGGCGTGACGGAGGGCCGCCGGCATGCGCCGCTCATGTTCGAGGAGTCCGACACGGGCAGGCGG
64    A  G  L  L  I  V  C  L  A  G  V  T  E  G  R  R  H  A  P  L  M  F  E  E  S  D  T  G  R  R

                       Splice
1090  TCCAACCGACCAGCGGTCACCGAATGCCAGTTTGGCAAAGTTTTGCGCGAATTGGGGTCCACCTGGTATGCGGATTTGGGTCCACCCTTC
94    S  N  R  P  A  V  T  E  [ C  Q  F  G  K  V  L  R  E  L  G  S  T  W  Y  A  D  L  G  P  P  F

1180  GGAGTTATGTACTGCATCAAGTGTGAATGTGTGGCGATACCCAAGAAGCGGCGCATCGTTGCACGCGTCCAGTGTCGCAATATCAAAAAC    ]CR1
124   G  V  M  Y  C  I  K  C  E  C  V  A  I  P  K  K  R  R  I  V  A  R  V  Q  C  R  N  I  K  N

1270  GAGTGCCCGCCGGCCAAATGCGATGATCCCATCTCTTGCCCCGGAAAATGCTGCAAGACCTGTCCCGGCGATCGAAACGATACGGATGTA
154   E  C  P  P  A  K  C  D  D  P  I  S  P  G  K  C  C  K  T  C  P  G  D  R ] N  D  T  D  V
                                                                              *  *  *

1360  GCCTTGGATGTGCCCGTGCCCAATGAAGAGGAAGAGCGCAACATGAAACATTACGCTGCGTTGCTAACGGGCCGCACCTCCTATTTCCTC
184   A  L  D  V  P  V  P  N  E  E  E  E  R  N  M  K  H  Y  A  A  L  L  T  G  R  T  S  Y  F  L

1450  AAGGGTGAGGAAATGAAGTCCATGTACACCACCTACAATCCGCAGAATGTGGTGGCCACCGCCCGTTTCCTGTTCCACAAGAAGAATCTA
214   K  G  E  E  M  K  S  M  Y  T  T  Y  N  P  Q  N  V  V  A  T  A  R  F  L  F  H  K  K  N  L

1540  TACTACTCCTTCTACACCTCATCGCGAATCGGTCGTCCGCGTGCCATTCAATTCGTTGATGATGCGGGTGTAATCCTGGAGGAGCATCAA
244   Y  Y  S  F  Y  T  S  S  R  I  G  R  P  R  A  I  Q  F  V  D  D  A  G  V  I  L  E  E  H  Q

1630  CTGGAGACCACCTTGGCGGGCACTCTCAGTGTCTATCAGAATGCCACGGGCAAGATCTGCGGTGTCTGGCGACGAGTTCCACGTGATTAC
274   L  E  T  T  L  A  G  T  L  S  V  Y  Q  N  A  T  G  K  I  C  G  V  W  R  R  V  P  R  D  Y
                                           *  *  *

1720  AAGCGCATCCTGCGCGACGATCGTCTCCATGTTGTCCTCCTCTGGGGCAACAAACAGCAGGCCGAGTTGGCTCTGGCCGGAAAGGTGGCC
304   K  R  I  L  R  D  D  R  L  H  V  V  L  L  W  G  N  K  Q  Q  A  E  L  A  L  A  G  K  V  A

1810  AAATACACGGCCCTGCAGACGGAGTTGTTCAGTTCGCTACTGGAGGCACCACTTCCCGATGGCAAAACGGATCCCCAGCTGGCCGGAGCC
334   K  Y  T  A  L  Q  T  E  L  F  S  S  L  L  E  A  P  L  P  D  G  K  T  D  P  Q  L  A  G  A

1900  GGTGGCACAGCGATCGTGTCCACCAGCAGCGGTGCCGCCTCATCGATGCATCTCACCCTGGTCTTCAATGGTGTCTTTGGTGCCGAGGAG
364   G  G  T  A  I  V  S  T  S  S  G  A  A  S  S  M  H  L  T  L  V  F  N  G  V  F  G  A  E  E

1990  TACGCCGATGCAGCACTGAGTGTGAAAATTGAGCTGGCAGAACGGAAGGAGGTGATCTTCGATGAGATTCCACGTGTGCGCAAACCCTCT
394   Y  A  D  A  A  L  S  V  K  I  E  L  A  E  R  K  E  V  I  F  D  E  I  P  R  V  R  K  P  S

2080  GCCGAGATCAATGTCCTGGAGCTGTCGTCGCCCATTTCCATACAGAATCTTCGACTGATGTCGCGTGGCAAACTCCTGCTGACCGTGGAG
424   A  E  I  N  V  L  E  L  S  S  P  I  S  I  Q  N  L  R  L  M  S  R  G  K  L  L  L  T  V  E

2170  TCCAAGAAGTACCCACATCTGCGCATCCAGGGACACATCGTGACCCGAGCCAGCTGCGAAATCTTCCAGACCCTGCTGGCGCCGCACAGT
454   S  K  K  Y  P  H  L  R  I  Q  G  H  I  V  T  R  A  S  C  E  I  F  Q  T  L  L  A  P  H  S   ]SR1

2260  GCCGAATCCTCGACCAAGAGCAGCGGTTTGGCGTGGGTCTACTTGAACACCGATGGATCTCTGGCCTACAACATCGAAACGGAGCACGTG
484   A  E  S  S  T  K  S  S  G  L  A  W  V  Y  L  N  T  D  G  S  L  A  Y  N  I  E  T  E  H  V

2350  AACACCCGGGATAGGCCCAACATCAGTTTGATTGAGGAGCAGGGCAAGCGGAAGGCCAAGCTGGAGGATCTGACGCCGAGCTTCAACTTC
514   N  T  R  D  R  P  N  I  S  L  I  E  E  Q  G  K  R  K  A  K  L  E  D  L  T  P  S  F  N  F
                                 *  *  *

2440  AACCAGGCCATTGGTAGTGTGGAGAAGTTGGGTCCCAAGGTCCTCGAGTCGCTGTATGCCGGCGAACTGGGCGTTAATGTGGCCACCGAG
544   N  Q  A  I  G  S  V  E  K  L  G  P  K  V  L  E  S  L  Y  A  G  E  L  G  V  N  V  A  T  E
```

FIGURE 6

2530 CATGAGACGAGCCTGATCCGTGGCCGCCTAGTGCCCCGTCCAGTGGCCGATGCTCGGGACTCGGCGGAACCCATTCTGCTGAAGCGACAG
574 H E T S L I R G R L V P R P V A D A R D S A E P I L L K R Q

2620 GAGCACACGGATGCACAGAATCCACATGCCGTCGGCATGGCCTGGATGTCCATTGACAACGAGTGCAATCTGCACTACGAGGTGACGCTC
604 E H T D A Q N P H A V G M A W M S I D N E C N L H Y E V T L

2710 AACGGTGTGCCCGCCCAGGATCTGCAGCTGTATCTGGAGGAGAAGCCCATCGAGGCGGATTGGAGCGCCAGTGACGAGGAAATTGCTCGAG
634 N G V P A Q D L Q L Y L E E K P I E A I G A P V T R K L L E   SR2

2800 GAATTCAACGGCTCCTATCTGGAAGGCTTCTTCCTCAGCATGCCATCCGCCGAACTGATCAAGCTGGAGATGAGCGTCTGCTATCTGGAG
664 E F N G S Y L E G F F L S M P S A E L I K L E M S V C Y L E

2890 GTCCATTCCAAGCACTCCAAACAGCTTCTGCTGCGCGGCAAACTGAAGAGCACCAAGGTGCCGGGTCACTGCTTCCCCGTCTATACGGAC
694 V H S K H S K Q L L L R G K L K S T K V P G H C F P V Y T D

2980 AACAATGTTCCCGTGCCTGGCGACCACAATGATAACCATTTGGTGAACGGAGAGACCAAGTGCTTTCACTCCGGACGCTTCTACAACGAA
724 N N V P V P G D H N D N H L V N G E T K [C F H S G R F Y N E]   SR3

3070 TCGGAGCAGTGGCGCAGTGCCCAGGATTCCTGTCAGATGTGCGCCCTGTTTGCGTGGCCAATCCAGTTGCGAGGTGATCAAGTGTCCGGCT
754 [S E Q W R S A Q D S C Q M C A C L R G Q S S C E V I K C P A]   CR2

3160 CTCAAGTGTCAAGTCCACGGAGCAACTGCTTCAGCGTGATGGTGAATGCTGTCCCAGCTGTGTGCCCAAGAAGGAGGCCGCCGACTATTCA
784 [L K C K S T E Q L L Q R D G E C C P S C V P K K] E A A D Y S

3250 GCGCAATCCTCGCCAGCCACCAATGCCACCGAT GTGCTGCAACAGCGACGCGGCTGCCGCCTGGGCGAGCAGTTCCATCCCGCCGGTGCC
814 A Q S S P A T N A T D L Q Q R R G [C R L G E Q F H P A G A]   CR3

3340 AGTTGGCATCCATTCCTGCCGCCCAATGGCTTCGATACCTGCACCACCTGCAGCTGCGATCCCCTGACCCTCGAGATTCGCTGTCCCCGG
844 [S W H P F L P P N G F D T C T T C S C D P L T L E I R C P R]

3430 CTCGTCTGCCCGCCGTTGCAGTGCAGCGAGAAGTGGGCCTATCGTCCAGACAAGAAGGCATGCTGCAAGATCTGTCCGGAGGGCAAGCAG
874 [L V C P P L Q C S E K L A Y R P D K K A C C K I C P E G K] Q

3520 AGCAGTTCCAATGGACACAAGACGACGCCGAACAATCCCAATGTGCTGCAGGATCAGGCCATGCAGCGATCGCCGAGTCACAGTGCCGAG
904 S S S N G H K T T P N N P N V L Q D Q A M Q R S P S H S A E

3610 GAGGTTCTGGCCAACGGCGGATGCAAGGTGGTCAACAAGGTGTACGAGAACGGCCAGGAGTGGCATCCGATCCTGATGTCCCACGGCGAG
934 E V L A N G G [C K V V N K V Y E N G Q E W H P I L M S H G E]

3700 CAGAAGTGCATCAAGTGCCGCTGCAAGGACTCCAAGGTGAACTGCGATGCCAAGCGCTGCTCCCGCTCCACGTGCCAGCAGCAGACACGC
964 [Q K C I K C R C K D S K V N C D A K R C S R S T C Q Q Q T R]   CR4

3790 GTGACCAGCAAACGGCGTCTGTTCGAGAAACCGGACGCAGCTGCTCCGGCCATCGATGAGTTCTGCTCCACCCAGTGCCGGAGATCGAGG
994 [V T S K R R L F E K P D A A A P A I D E F C S T Q C R R S R]

3880 CGCCACCACAAGAGGCAGCCGCATCATCAGCAGCGATCCTCCAGCTGAGCGGCTCCACGTGACGGATGGGATCCCAATCCAGTATCAGAT
1024 R H H K R Q P H H Q Q R S S S end

3970 CCTTGGCGGCAGGGGAGCGAACCAATCACTCACTCACTCACCACCACTCAGTGTACTCAGTGTGCACCACCCAAACACACACACACACAC
4060 ACACACACAACCACACAACACTCACACCCACATCTACACAGACACACAGACAGCCACAAAAGCGAACGCGCACACAGACTTGTGCAAGGA
4150 GTTGCATAGATCGTTGTTGCTATCTTATCATGTGGCAGCAATGAGAACTTGTATTATATATATGAATCACGGAGGAGAAAACGTAGGAGA
4240 GAAATCTCACAAAAAATATATATATCTTATGGAGGAAAACGGTAGTAATAGAGAGAGAGAGAGAGGGAAGGAGAGAGTCTAATGAGATCC
4330 TTGGAAAAGGACATTAAAACCAGTGCAGTTTGCTTTAAATTCTCCAGCGCAGAATTTTCTATTGAAAGCATTTTCTGAATTTCTTTTCGC
4420 AGTTACCCCACCCGTGTAACCCAATCCCCTCCCCTCCCCAACCAACAAACACCCAAAAAAAAAAAACTAAAAACATTAAAATACAATTTTA
4510 ATTTATTACAAAACAAAAAACAAAAAAAAAAAA

FIGURE 6

sub C

TM

CR1

FIGURE 1

FIGURE 2

FIGURE 3

FIGURE 4

Sequence Range: 1 to 232

```
            10        20        30        40        50        60        70
Nog protei  MERSQCLVTTRLMVTLGLRDQGGCQHTLHTRAPSELPLVDLIHPDPTDPKEQLNETLRTLMVGHTP
```

```
                                                    260       270       280
SuperSog P                                          iqivddaqvile--EhqLetTLagTLsV--yqr
[   71 ]                                             ^^^^^^^^^~~~   ^~~~~_~~^y~~~   ^^y
Nog protei                                          VDLIHPDPTDPKEQLNETLRTLMVGHTP
```

```
            90        100       110       120       130       140       150
Nog protei  ILPERLGVEDLGEEDLLRQPSGAMPAEIKGLEFYEGLQSGHPLSHTLPRLCQMWLWSQTFCFVLYTWNDLC
```

```
            |        300       310
SuperSog P  |        IergsRvplEELcEgtLLLw>
[   71 ]    |        ^v----^^^^^^^^^y-^^^^
Nog protei  |  ILPERLGVEDLGEEDLLR
```
pris vector

```
            170       180       190       200       210       220
Nog protei  RVVRVGSCYSRSCSVPEGMVCKRAKSMELTILPWRCQRVQQCQWITIQYFVISECKCSC
```

FIGURE 5

sub C

TM

CR1

SR1

SR2

SR3

CR2

CR3

CR4

FIGURE 6